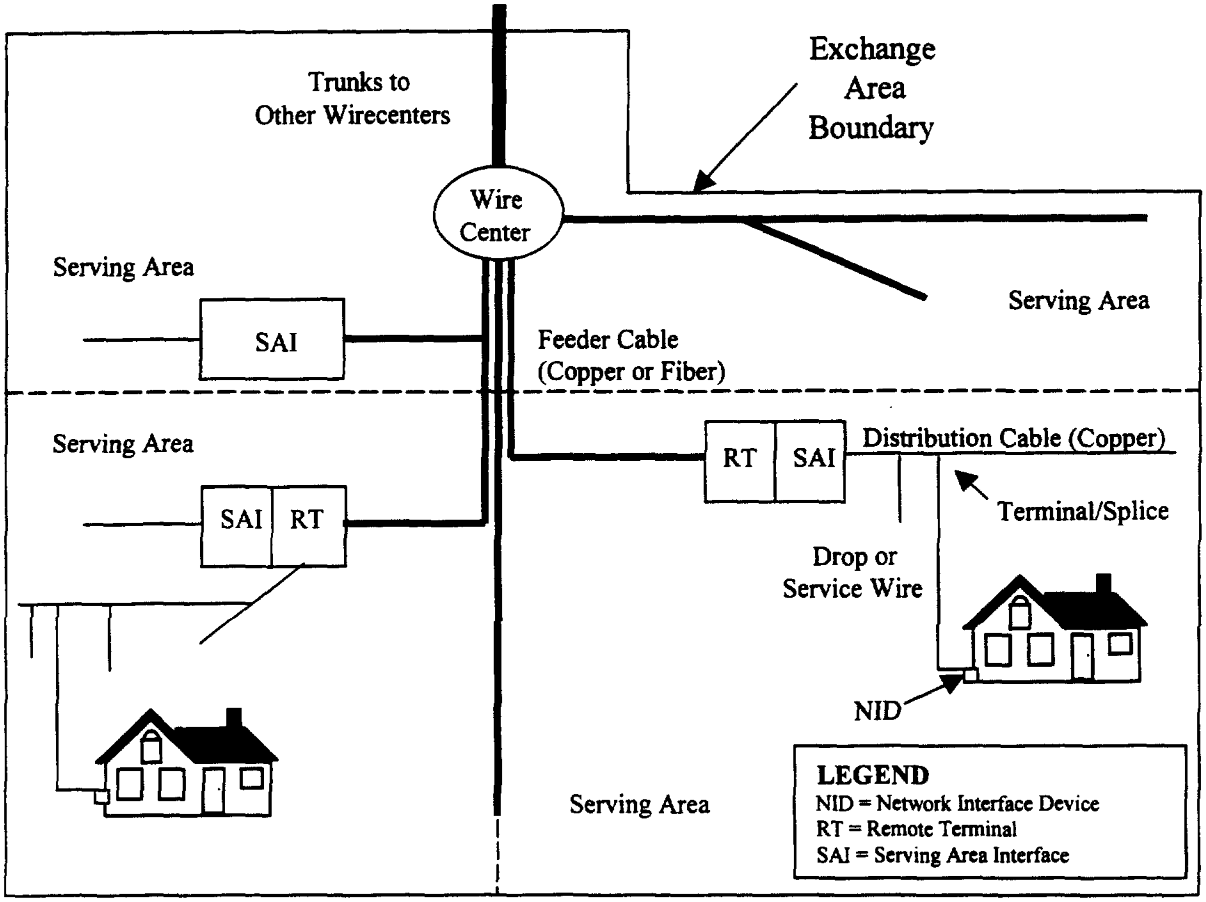

Adapted from Engineering and Operations in the Bell System, 2nd Edition, 1983

Figure 1 Loop Components

3.1.1.1. Serving Area

The total area served by a wire center is organized into one or more serving areas, each of which contains a portion of the area and lines served by the wire center. The serving areas are delineated by dotted lines in the above figure. In HM 5.0, the serving areas equate to main customer clusters and their subtending outlier clusters, as discussed in Section 6.2.

3.1.1.2. General Loop Description

One end of the feeder portion of the loop terminates within the central office building, or "wire center." Copper cable feeder facilities terminate on the "vertical side" of the main distributing frame ("MDF") in the wire center, and fiber optic feeder cable serving integrated digital loop carrier ("IDLC") systems terminates on a fiber distribution frame in the wire center.

The other end of the feeder extends to an appropriate termination point in the serving area. Copper feeder cables terminate on one or more serving area interfaces ("SAIs") in each serving area, where they are cross-connected to copper distribution cables. Fiber feeder cables extend to a digital loop carrier ("DLC") remote terminal ("RT") in the serving area, where optical digital signals are demultiplexed and converted to analog signals. Individual circuits from the DLC are cross-connected to copper distribution cables at an adjacent SAI.

Copper distribution cable extends from the SAI along routes passing individual customer premises. At appropriate points, these cables pass through block terminals typically serving several housing units. In the terminal, individual copper pairs in the distribution cable are spliced to "drops" that extend from the terminal to the customers premises. The drop terminates at a network interface device, or NID, at the customer's premises.

Feeder, distribution, and drop cables are supported by "structures." These structures may be underground conduit, poles, or trenches for buried cable and underground conduit. Underground cable is distinguished from buried cable in that underground cable is placed in conduit, while buried cable comes into direct contact with soil.[12] In more urban areas, aerial distribution cable may be attached directly to the outside of buildings, in what is called a "block cable" arrangement, or, for high-rise buildings, may consist of riser cable inside the building.

3.1.1.3. Local Loop Components

1) *Network Interface Device*

   The NID is the demarcation point between the local carrier's network and the customer's inside wiring. This device terminates the drop wire and is an access point that may be used to isolate trouble between the carrier's network and the customer's premises wiring. The NID also contains protection against externally-induced hazardous voltages, such as those associated with lightning strikes and contact between telephone and electric lines. In a multi-tenant building, the protection is located at the point at which the distribution cable enters the building.

2) *Drop*

   A copper drop cable, typically containing several wire pairs, extends from the NID at the customer's premises to the block terminal at the distribution cable that runs along the street or the lot line. The drop can be aerial or buried; generally it is aerial if the distribution cable is aerial, and buried if the distribution cable is buried or underground.

3) *Block Terminal*

   The " block terminal" is the interface between the drop and the distribution cable. When aerial distribution cable is used, the block terminal is attached to a pole in

[12] Although the conduit supporting underground cable is always placed in a trench, buried cable may either be placed in a trench or be directly plowed into the earth.

the subscriber's backyard or at the edge of a road. A pedestal contains the block terminal when distribution cable is buried.

4) *Distribution Cable*

Distribution cable runs between the block terminals and an SAI located in the serving area. Limitations on the capacity of an SAI and/or the distribution design used in a particular serving area may lead to multiple SAIs. Distribution structure components may consist of poles, trenches and conduit.[13]

5) *Conduit and Feeder Facilities*

Feeder facilities constitute the transmission system between the SAI and the wire center. These facilities may consist of either pairs of copper wire or a DLC system that uses optical fiber cables as the transmission medium. In a DLC system, the analog signals for multiple individual lines are converted to a digital format and multiplexed into a composite digital bit stream for transmission over the feeder facilities

Feeder structure components include poles, trenches and conduit. Manholes for copper feeder or pullboxes for fiber feeder are also normally installed in conjunction with underground feeder cable. Manhole spacing is a function of population density and the type of feeder cable used. Pullboxes that are installed for underground fiber cable are normally farther apart than manholes used with copper cables, because the lightness and flexibility of fiber cable permits it to be pulled over longer distances than copper cable.

Several utilities, e.g., electric utilities, LECs, IXCs and cable television ("CATV") operators, typically share structure because it is economical to do so. Manholes may be shared with low-voltage facilities. The amount of sharing of structure and manholes may differ in different density zones and between feeder and distribution cables.

3.1.2. Switching and Interoffice Network Description

This section provides a general description of the network components comprising the wire center and interoffice facilities. Figures 2 and 3 illustrate the relationships among the components described below.

3.1.2.1. Wire Centers

[13] Because underground distribution exists only in the highest density zones where runs are relatively short, and because in such zones underground structure is commonly shared with feeder, distribution facilities typically do not include manholes.

The wire center is a location from which feeder routes extend towards customer premises and from which interoffice circuits or "trunks" emanate toward other wire centers. A wire center normally contains at least one end office ("EO") switch and may also contain a tandem switch, an STP, an operator tandem, or some combination of these facilities. Wire center physical facilities include a building, power and air conditioning systems, rooms housing different switches, transmission equipment, distributing frames and entrance vaults for interoffice and loop feeder cables.

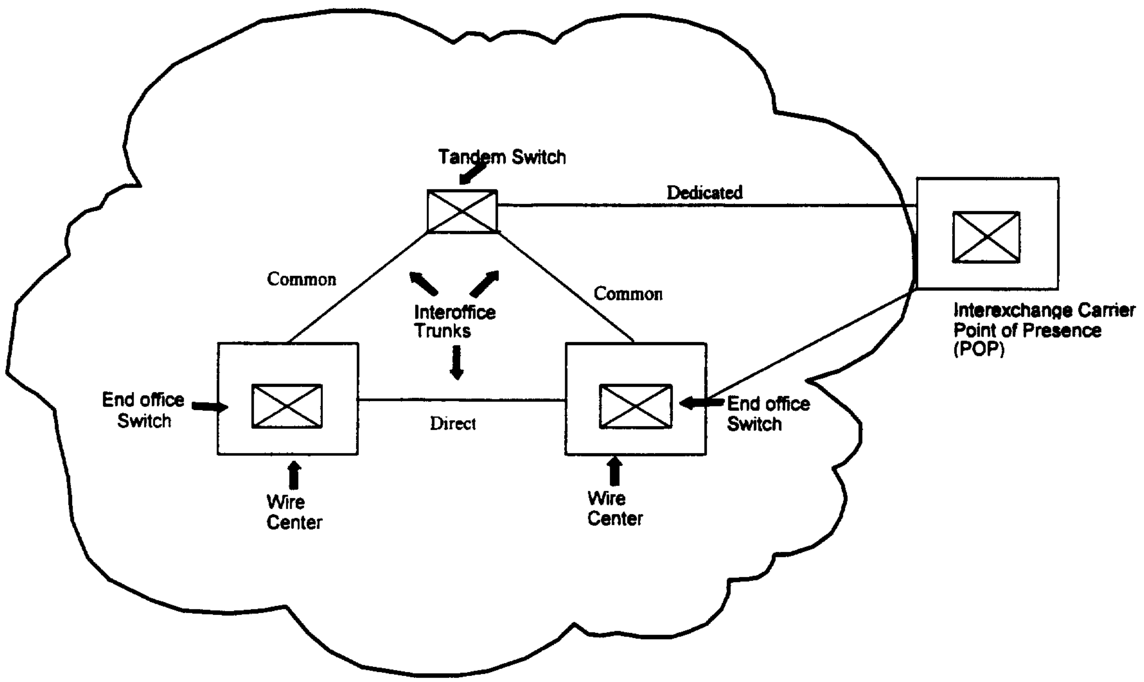

Figure 2 Interoffice Network

3.1.2.2. End Office Switches

The end office switch provides dial tone to the switched access lines it serves. It also provides on-demand connections to other end offices via direct trunks, to tandem switches via common trunks, to interexchange carrier ("IXC") points of presence ("POPs") via dedicated trunks, and to operator tandems via operator trunks.

3.1.2.3. Tandem Switches

Tandem switches interconnect end office switches via common trunks, and may also provide connections to IXC POPs via dedicated trunks. Common trunks also provide alternatives to direct routes for traffic between end offices. Tandem switching functions often are performed by switches that also perform end office functions. Tandems normally are located in wire centers that also house end office switches

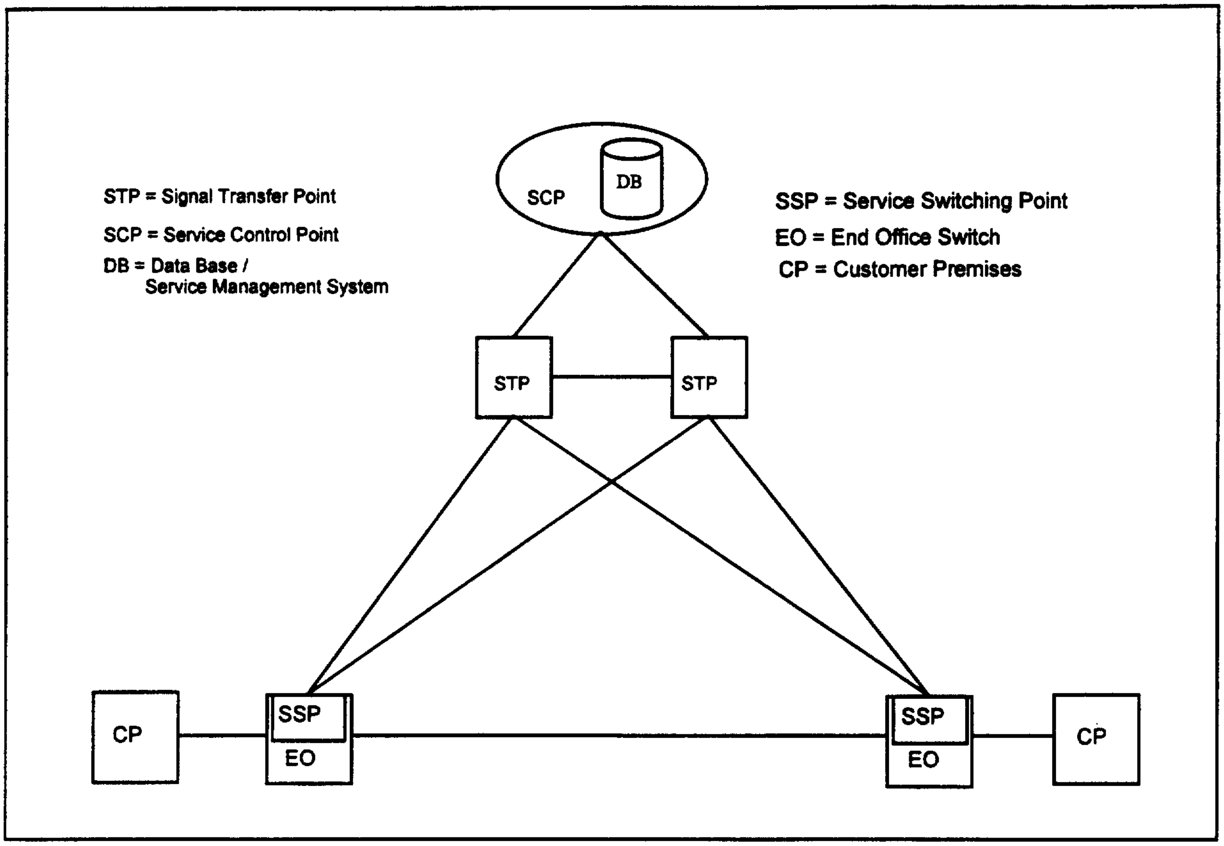

Figure 3 Interoffice Signaling Network Components

3.1.2.4. Interoffice Transmission Facilities

Interoffice transmission facilities carry the trunks that connect end offices to each other and to tandem switches. The signaling links in a Signaling System 7 ("SS7") signaling network are also normally carried over these interoffice facilities.

Interoffice transmission facilities are predominantly optical fiber systems that carry signals in SONET format. Both economic and service quality considerations increasingly prescribe the use of a fiber optic ring configuration to link switches, except for switches that serve few lines or that are too remote from other switches, where ring costs might be prohibitive. In this case, the small switches are typically connected to the tandem via point-point links that are increasingly provided on a route-diverse (that is, redundant) basis for the sake of increasing reliability. Use of either rings or the redundant point-to-point links provides a redundant path between any two switches, and the potential for substantial cost savings relative to the ubiquitous deployment of traditional point-to-point facilities

3.1.2.5. Signal Transfer Points

STPs route signaling messages between switching and control entities in a SS7 network. Signaling links connect STPs and Service Switching Points ("SSPs"). STPs are equipped in mated pairs, with at least one pair in each Local Access Transport Area ("LATA").

3.1.2.6. Service Switching Points and Signaling Links

SSPs are SS7-compatible end office or tandem switches. They communicate with each other and with Service Control Points ("SCPs") through signaling links, which are 56 kbps dedicated circuits connecting SSPs with the mated STP pair serving the LATA.

3.1.2.7. Service Control Points

SCPs are databases residing in an SS7 network that contain various types of information, such as IXC identification or routing instructions for 800 numbers in regional 800 databases, or customer line information in Line Information Databases ("LIDB").

4. HM 5.0 Model Organization, Structure and Logic

4.1. Overview of HM 5.0 Organization

Figure 4 shows the relationships among the various modules contained within HM 5.0. An overview of each component of the Model follows.

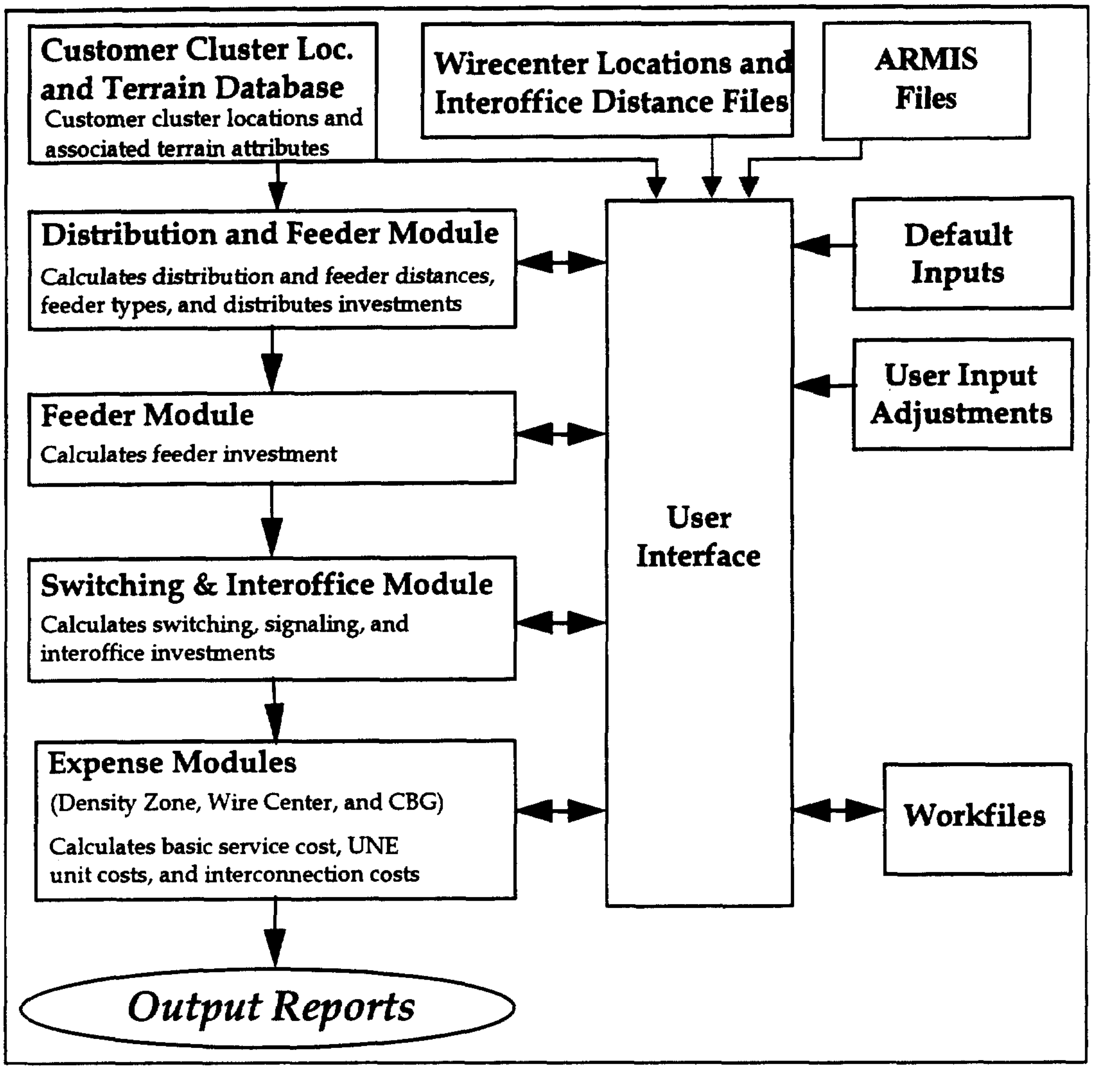

Figure 4 HM 5.0 Organization Flow Chart

4.2. Input Data

Inputs to HM 5.0 include detailed data describing the following items.

- Demographic, geographic and geological characteristics of the areas served by each telephone company. These data records are specific to individual "clusters" of customer locations (i.e., "main" clusters and their subtending "outlier" clusters). They include information on the number of customer locations and lines receiving telephone service in a cluster, the geometric configuration of the cluster's boundaries, the wire center that serves the cluster and the cluster's location relative to this wire center and to other nearby clusters, and the type of terrain that characterizes the cluster. The character and development of these data are described in greater detail in Section 5.

- Wire center locations, and interoffice distances between end offices, tandems, and STPs used to determine the route miles required for interoffice transmission and signaling facilities. These data are largely developed from Bellcore's LERG and NECA Tariff 4.

- 1996 ARMIS data reported by the Tier 1 LECs. These data provide information about current demand levels that the LEC must serve, and relationships between the LEC's embedded expenses and investments.

- Numerous user-adjustable inputs that allow users to set carrier- or locale-specific parameters, and perform sensitivity analyses. These inputs have preset default values based on the engineering experience and judgment of HAI personnel, as well as the judgements of independent subject matter expert consultants to HAI.

4.3. Workfiles

Workfiles are created from the input data files when a particular state/company (study area) combination is run for the first time. As a complete run of the HM 5.0 progresses, intermediate outputs from the HM 5.0's constituent modules are stored in these workfiles. Once the run is complete, the workfile may be examined. A great deal of information above and beyond that presented in the Expense Module spreadsheets (that contain the principal final results of the model's analysis) may be obtained from these workfiles. One example is average loop length, by cluster and by wire center. Additionally, the user may perform further analyses on the LEC study area by directing the model to create a new workfile.

4.4. User Interface

The Hatfield Model includes a user interface program that facilitates model operation, including extraction of data from the input files, providing dialog boxes for users to manipulate model inputs, executing the Excel workbooks that constitute the model, saving intermediate results, and managing the flow of intermediate results between different modules.

The user interface program also performs certain simple aggregation and summarization calculations in Visual Basic for Applications ("VBA"). This shortens greatly execution times and allows users examining the model's Excel workbooks to focus on the model's fundamental engineering logic.[14]

4.5. Distribution Module

[14] Model versions prior to HM 3.0 used Microsoft Excel's Pivot Table feature to summarize various results at the wire center and density zone levels. Although this feature was quite flexible, applying pivot tables to the very large arrays of data required by the model led to very slow execution times.

The Distribution Module addresses the portion of the network extending from SAIs to the customers' premises. The module determines the lengths and sizes of distribution cable, the associated structures (poles, trenching, and conduit), the number of terminals, splices, drops, and NIDs required to provide service to the specified numbers and types of customers, and the number and type of SAIs and DLC terminals required. The module also calculates certain distances required by the Feeder Module, and, according to those calculations, determines whether to serve a given distribution area using feeder transmission facilities consisting of copper wire pairs, or using DLC running over optical fiber cable. The model selects fiber feeder if any of following five criteria are met:

a) the feeder distance exceeds a user-adjustable crossover distance (set to a default value of 9,000 feet) that limits maximum distance of any copper feeder run;

b) the total analog copper loop within a main cluster, including feeder and distribution, exceeds a user-adjustable distance whose default value is 18,000 feet;[15]

c) the main cluster has at least one outlier cluster subtending it;

d) an analysis of the life-cycle costs of fiber vs. copper feeder shows that fiber feeder is the more economical choice, or

e) the "wireless" investment cap is invoked.

These criteria are described in greater detail in Section 6.3.5. If, based on these criteria, copper feeder is chosen, it extends out to an SAI located at the centroid of the main cluster. Copper distribution cable then extends from the SAI to all customer premises in the cluster. If fiber feeder is chosen, it extends from the wire center out to a DLC RT located at the centroid of the main cluster, and copper distribution cable then extends from the DLC RT to all customer premises in the main cluster.[16]

The HM 5.0 Distribution Module serves outlier clusters that subtend main clusters with analog copper cable if their distance from the DLC RT in the main cluster does not exceed a (default 18,000 feet) user-adjustable distance parameter, and if this outlier cluster has no other outlier clusters either subtending it, or lying between it and the main cluster.[17] If the distance to the farthest subscriber within an outlier cluster would exceed this threshold, the Distribution Module serves the outlier cluster with digital loop carrier equipment using copper-based T1 transmission.[18] Once the outlier cluster has been

[15] One criterion that defines any cluster is that no customer location may be more than 18,000 feet from the centroid of the cluster along a right-angle route.

[16] If, however, the model invokes a wireless local distribution system, fiber feeder extends to the radio base station located in main cluster, and final distribution is made to customers in outlier clusters over-the-air.

[17] Such an outlier cluster is termed a "first order" outlier.

[18] This technology does not require the use of loading coils or coarse-gauge cable, and it also permits basic rate ISDN and other advanced narrowband services to be provided to all subscriber locations in the model.

reached, analog copper distribution cables are used to serve the customers located in the outlier cluster.

After the module has determined the quantities of all distribution elements, it calculates the investment associated with these elements, including distribution and drop cable, structures, NIDs, terminals and splices, SAIs, and DLC terminals, using as inputs the user-adjustable unit prices of each element. The numbers and types of elements engineered can be examined in the intermediate outputs of the Distribution Module as recorded in the workfile.

4.6. Feeder Module

The Feeder Module configures the portion of the network that extends from the wire center to the SAIs. Based on information it receives from the Distribution Module, it determines the size and type of cables required to reach the SAIs located in each serving area, along with supporting structures (poles, trenching, conduit, manholes, and fiber optics pullboxes). The Feeder Module then calculates the investment associated with these elements, using as inputs the unit prices of each such element. It passes these investments to the Expense Modules. The numbers and types of network elements required can be examined in the intermediate outputs of the Feeder Module as recorded in the workfile.

4.7. Switching and Interoffice Module

The Switching and Interoffice Module computes investments for end office switching, tandem switching, signaling, and interoffice transmission facilities. In HM 5.0, the user can designate specific wire center locations that house host, remote, and stand-alone switches, respectively, as well as specify inputs for the per-line investments associated with each type of switch. HM 5.0 will then calculate switching investments taking into account the switch arrangements that the user designates.

The switching module determines the required line, traffic, and call processing capacity of switches based on line totals by customer type across all serving areas belonging to the wire center, and based on ARMIS-derived traffic and calling volume inputs. It also determines the required capacity and distances of interoffice transmission facilities, using the traffic data and the interoffice distances that are input to the Module. In doing so, it uses wire center locations and interoffice distances to determine an efficient mix of interoffice SONET rings and redundant point-to-point links (for small offices not exceeding a user-adjustable line size, default 5000 total lines). Rings are separately provided for linking host switches to their subtending remotes, and for linking host switches to each other, to stand-alone switches and to the tandem switches on which they home. Even when the model determines a given office should be directly linked to its tandem, it attempts to physically route these point-to-point circuits through a nearby large wire center and then over the fiber rings it has otherwise engineered for interconnecting larger offices and tandems The numbers and types of elements involved can be examined in the intermediate outputs of the Switching and Interoffice Module as recorded in the workfile.

4.8. Expense Modules

There are three different versions of the HM 5.0's Expense Module – one each for the three levels of granularity by which the user can elect to have cost results displayed: by line density range, by wire center, or by CBG.[19] Each version calculates the monthly costs for unbundled network elements, universal service, and carrier access and network interconnection. These costs include both the capital carrying costs associated with the investments, and the costs of operating the network. Capital carrying costs include depreciation, return on the debt and equity investment required to build the network, and income taxes on equity return. Network-related expenses include maintenance and network operations. Non-network related expenses include customer operations expenses, general support expenses, other taxes and variable overhead expenses.

Several sources provide information to the Expense Modules. The Distribution, Feeder, and Switching and Interoffice Modules provide network investments by specific plant category. ARMIS and other sources are used to derive information on network operating and maintenance expense relationships.

The Expense Modules produce reports (either by density zone, wire center or CBG) showing the key outputs of the model, including the costs of providing universal service, unbundled network elements, interconnection and IXC access. Further detail about network investments and costs is available from the workfile associated with a model run.

5. Input Data

To accommodate HM 5.0's evolution to modeling local telephone networks based on actual clusters of customer locations, the input data used in HM 5.0 are much more granular than the CBG input data used in HM 4.0 and earlier. Flowcharts describing the development processes used to prepare these input data for HM 5.0 are attached as Appendix C to this document.

5.1. Line Type Counts by Study Area

[19] Although the HM 5.0 engineers no plant based on CBG granularity, the results of its engineering to individual clusters may be rolled up to display cost results at the CBG level. The association between clusters and CBGs is made based on the CBG that accounts for the plurality of the cluster's land area.

Counts of access lines by type (i.e., residence, single line business, multiline business, public telephone and special access lines) for each distinct NECA Study Area for calendar year 1996 are developed from several data sources. These include:

ARMIS 43-08, 10/01/97;

ARMIS 43-01, 10/01/97;

NECA USF Loops filing, 1996;

USTA report, 1995;

RUS report, 1995;

1993 USF Data Request;

ARMIS-based line factors.

The rules by which the best of these data are selected are as follows.

a) When NECA Study Area name matches exactly ARMIS Company name, populate line types directly from ARMIS data for business lines (43-08), single line business lines (43-01), residence lines (43-08), special access lines (43-08), and public lines (43-08) data.

b) For remaining ARMIS Companies, determine counts of line types for NECA Study Area name by applying ARMIS line type distributions to total reported NECA USF Loops.

c) For non-ARMIS Companies, match NECA Study Area name to best available data source (i.e., USTA, RUS or USF Data Request) for residence and business line splits.

d) When no company-specific line type data exist, apply average ARMIS line type distributions to reported NECA USF Loops for NECA Study Area name.

5.2. *Wire Center List*

The source of the wire center information used in PNR's National Access Line Model is Bellcore's LERG database, dated August 1, 1997.[20] The Hatfield Model developers have licensed from Bellcore an extract of key data from the LERG called the Special LERG Extract Data ("SLED").

Certain switching entities (wire centers) in the SLED with Common Language Location Identifier ("CLLI™") codes not marked as end offices, hosts or remotes are then removed from this wire center database. In addition, switching entities that are inactive, or owned by wireless, long distance or competitive access providers are removed as well.

[20] These LERG data are augmented by data from NECA Tariff 4.

In a few instances, the SLED assigns wire centers to multiple local carriers. This may result from switch collocation. Because the Hatfield Model requires wire center entries to be unique, such wire centers are assigned to the local carrier having the greatest number of active NPA-NXX codes. If active NPA-NXX codes are equal among companies, assignment is to the carrier having the greatest number of residential lines.

Multiple occurrences of 8-character CLLIs may also occur in the SLED due to placements of several switches at a single wire center location. Because the Hatfield Model itself engineers multiple switches in a wire center if demand requires it, duplicate occurrences of 8-character CLLIs are removed from the model's wire center list.

5.3. Customer Counts by Census Block and Wire Center

Customer locations must be associated both with CBs, as well as their serving wire center ("WC"). The PNR National Access Line Model, Version 2.0 ("NALM") performs both of these tasks. The PNR NALM uses PNR survey information, Bellcore's LERG, BLR wire center boundaries, Dun & Bradstreet's ("D&B") business database, Metromail's household database, Claritas' 1996 demographic database, and U.S. Census estimates to calculate both the number of residential and business locations and access lines in each CB, and in each wire center in the United States. This summary describes the methodology, data and assumptions used in developing these location and line estimates in the NALM.

5.3.1. Residence Counts

Residential customer location counts are developed by applying the following process.

a) The Metromail household database (described in section 5.4.1, below) is geocoded to the "point" level.[21] In addition to recording the precise six-decimal place latitude and longitude of this household, the CB associated with its location is recorded as well. Duplicate household information is identified and eliminated. If two records appear with an identical latitude, longitude and phone number, one of the two records is eliminated.

b) Implied residential household counts are evaluated by comparing Metromail counts to Claritas' 1996 CBG-level projections of households with telephones. When Metromail households exceed Claritas households, Metromail households are used. When Claritas households exceed Metromail households, Claritas households are used, and the total differences are distributed to the constituent CBs in proportion to 1990 U.S. Census household distributions.

c) Access line counts are determined from household counts using probabilities, that is, how likely is it that a household will have a first or second telephone line

[21] As described in more detail in Section 5.4.3, below, geocoding to the "point" level means that the geocoding software has both found the housing unit's address in its location files and determined a latitude and longitude for the location down to six decimal places of a degree.

installed? First line probabilities are provided by Claritas based on demographic age and income profiles by CBG. Second line probabilities are based on a logistic regression using similar demographic information and developed by PNR using its ReQuest™ III residential survey. Multiplicative probability factors are applied to the household counts defined above to derive residential line counts.

d) The above derived residential line counts by CB are then normalized to sum to Study Area wide data on total residential line counts developed in Section 5.1, above.[22]

e) This lines normalization factor is applied to the residential customer location counts in each CB, as well.

The implications of the forgoing process are as follows. Because the primary source of residential location counts is Metromail – which includes all residences that receive direct mail regardless of whether they have telephones or not – the universe of "populated" CBs that the data process captures may include CBs where telephone service is not currently offered or accepted. Thus, the "breadth" of the telephone network that these data will instruct the HM 5.0 to construct is likely greater than the embedded networks of the ILECs. However, because the counts of lines and locations in each CB are normalized to sum to given Study Area wide totals, the "depth" of the constructed network will be consistent with current levels of actual telephone demand.[23]

5.3.2. Business Counts

Business location counts are developed by applying the following process.

a) The D&B national business database (described in Section 5.4.2, below) is geocoded to the "point" level. In addition to recording the precise six-decimal place latitude and longitude of this business, the CB associated with its location is recorded as well.

b) From the D&B national database, the total number of business lines, as well the probabilities of these lines being single line business lines and multiline business lines such as Centrex and PBX lines are developed. This model is based on an 800,000 firm sample.

c) Because the D&B national business database contains records for only about 11 million out an estimated total of 12 million U.S. businesses, and because the

[22] If comprehensive LEC data on residential line counts by individual wire center are available, normalization can be done at the individual wire center level.

[23] In addition, note that these primary source residential location counts derive from precisely geocoded 1997 Metromail data. Thus, these data provide 1997 information on location counts at the CB level. As a result, the model's reliance on noncurrent or nonCB-specific location data (e.g., Claritas 1996 CBG-level projections, 1990 U.S. Census CB counts, or 1995 U.S. Census Update county-level projections) is limited to those locations that show up in such counts that are in excess of the Metromail counts.

businesses that it misses are almost certainly small businesses, an additional 1 million nonD&B business locations are added to CB counts in proportion to D&B businesses located in the CB. The lines associated with these added business locations are projected by PNR based on an assumption that they employ, on average, between 1 and 4 employees, each.

d) The above derived business line counts by CB are then normalized to sum to Study Area wide data on total business line counts developed in section 2.1, above.[24]

5.3.3. Location and Line Counts by Wire Center

HM 5.0 uses WC boundaries provided by BLR as its primary source to define wire center service areas. These boundaries conform to CB boundaries, and customer locations contained within all of the CBs associated with a WC are then assumed to be served by that WC. Telephone number information (NPA-NXX) continues to be used for backup and data scrubbing purposes when anomalies arise in the BLR geographical assignment process – as can occur if one wire center's boundaries fall completely within another wire center's boundaries.

5.4. Customer Location

The customer location approach used in HM 5.0 is fundamentally different from that of HM 4.0 – or any other approach that uses arbitrary geographic delineators such as CBs, CBGs or latitude and longitude grid cells. Because HM 5.0's approach identifies the actual locations (accurate to within 50 feet) of most telephone customers, it produces the most sophisticated demographic data set of its type. The process first develops a database of more than 110 million customer address records. These addresses are then geocoded (assigned latitude and longitude coordinates). These locations are then divided among wire center serving areas based on geocoded customer location and the BLR wire center boundaries.

5.4.1. Residence Location Data

Data for residence locations are provided by Metromail, Inc. The Metromail National Consumer Database© ("NCDB") is a large, nationally compiled file of U.S. household-level consumer information that includes both deliverable postal addresses (and telephone numbers, when available). The file consists of over 100 million households – which constitute over 90% of the 110 million households that the U.S. Bureau of the Census reported for 1995.

To ensure that the data captured are the most current available, this file is updated 65 times per year, and undergoes numerous "hygiene" measures to ensure its continued high

[24] Again, if comprehensive LEC data on business line counts by individual wire center are available, normalization can be done at the individual wire center level.

quality for direct marketing purposes. Such purposes require the data to reflect postal address standardization practices, incorporate National Change of Address ("NCOA") processing, and permit postal geocoding to street address, ZIP+4 or Carrier Route levels.

The file is compiled primarily from telephone white pages directory data, but also utilizes many other primary sources of information, such as household mover records, voter registration data, motor vehicle registration information, mail-order respondent records, realty data, and home sales and mortgage transaction information, to build a large repository of verified household-level data.

5.4.2. Business Location Data

Dun & Bradstreet collects information on more than 11 million business establishments nationwide. Information is gathered from numerous sources such as business principals, public records, industry trade tapes, associations, directories, government records, news sources, trade organizations, and financial institutions. This information is validated each night. Additionally, D&B conducts millions of annual management interviews to help improve the timeliness and accuracy of its information.

The information is organized by D-U-N-S number, a nine digit identification sequence which allows for the placement of companies within larger business entities according to corporate structures and financial relationships. A D&B family tree may be used relate separate operating companies to each other, and to their ultimate parent company. D&B also provides "demographic" information on each of the firms in its database. Such information includes counts of employees and the SIC code of the establishment.

5.4.3. Geocoding

Geocoding is used in order to most accurately assign known customer locations to actual, physical locations. Geocoding is also known as location coding. It involves the assignment of latitude and longitude coordinates to actual street addresses. Geocoding software is sophisticated enough to provide information regarding the source and precision of the lat/long coordinates selected. This precision indicator allows PNR to select only those addresses that have been geocoded to a highly precise point location. Geocoding also allows customer location points to be assigned less granularly to the CBG level, or higher. Almost uniformly, geographical address locations are derived from enhanced versions of the USGS' TIGER database.

To perform its geocoding, PNR uses a program by Qualitative Marketing Software called Centrus™ Desktop. The enhanced data behind Centrus is provided by GDT. Premium GDT data are updated bi-monthly to ensure accuracy. These data integrate new information from US Postal Service ("USPS") databases and private sources so that new streets and additions and changes to ZIP codes, street names, and address ranges are included as soon as possible.

Centrus™ Desktop allows geocoding on two levels. The first is a match to the actual address -- which is the only type of geocoding used in HM 5.0 customer location. The second is a match to a ZIP code (ZIP, ZIP+4, ZIP+2) level. Because of the lesser accuracy in the second method, these geocodes are not used in PNR's process of assigning customer locations.

Within the geocode process, there are a number of options available to the user. Each of these options determine the quality of the matches allowed in the end-use geocode. For purposes of customer location, addresses are always matched to the "Close" setting. "Close" allows for minor misspellings in addition to incorrect or missing directionals (North, East, etc.) or street types (street, road, etc.). Although ZIP-based geocodes are generally accurate enough for most applications, they are not considered good enough for actual customer locations and are not used to develop and locate customer clusters in HM 5.0.

Data hierarchy in address geocoding starts with the State. The hierarchy continues with City, Street Name, Street Block, and finally, House Range. Typically, a Street Block is the same as an actual physical block but it can also represent a partial block as well. The House Range displays address information from the USPS. Additionally, where there are gaps in the actual address range, the House range will account for these gaps.

Initially, the address coding module in Centrus™ Desktop compares the street addresses from the input file to the records contained in the USPS ZIP+4 directory and the enhanced street network files. If the address is located in the USPS files, the address is standardized and a ZIP+4 is also returned. If this address is also found in the street network files, Centrus™ Desktop determines a latitude and longitude for the location. Optionally, if the address is not found in the street network files, location information may be applied from the ZIP level.[25]

Location codes generated by Centrus™ Desktop indicate the accuracy of the geocode. For purposes of customer location clustering in the HM 5.0, only those geocodes assigned at the 6-decimal place point location made directly to the street segment are used.[26]

While the software and data used allow for a much more comprehensive output of data elements, for use in HM 5.0 customer location, the following addressing elements are extracted:

Address

---

[25] Note that ZIP+4 codes may be very precise. In general, they are specific to the face of single city block. While it may turn out that accuracy to the street block face is quite sufficient for accurate cost modeling of local telephone networks, in the interest of conservatism, these type of geocodes are not presently used in HM 5.0 data.

[26] Furthermore, placement of the address along the street segment is quite precise. The Centrus geocoding software and reference data also make use of USPS determinations of whether the segment contains a continuous or discontinuous range of address numbers. Thus, if the addresses on a block face run from 200 to 250 and 274 to 298 (with the range between 252 and 272 missing), an address of 250 will be geocoded, it will not simple be geocoded as at midblock.

City
State
ZIP
ZIP+4
Latitude
Longitude
Census Block
Match Code
Location Code

5.4.4. Gross-up

The above-derived precisely geocoded locations are then counted by CB. These geocoded location counts by CB are then compared to target total line counts for that CB derived by the PNR NALM (as described in section 2.3, above). If the geocoded location counts are less than the target count, the residual number of customer location points is then computed, and geographical locations for these points are generated. This process is performed by PNR using TIGER file CB boundaries. Each of the additional number of customer location points that a CB requires to total to its target count is generated and assigned a geocode so as to place these "surrogate" points uniformly along the CB's boundary. While these boundary-assumed locations for the gross-up or surrogate points are plausible – because most CBs are bounded by roads – this is also the most conservative placement of the gross-up points because it assumes they are maximally separated from one another.

As a result of this gross up process, the customer location file now contains records for each of the U.S.'s more than 100 million customer locations with a geocode (either calculated precisely or through the gross up process) associated with it.

5.5. Customer Location Clustering

5.5.1. General Criteria

The input development process next identifies all customer locations within a wire center's boundaries that are close enough together to be efficiently engineered as a single telephone plant serving area. This process is called clustering. While there are many available off-the-shelf clustering algorithms, efficient determination clusters of customer locations that are consistent with telephone engineering practices requires that certain engineering restrictions be imposed during the clustering process, and not afterward. Customer locations must meet the following criteria to be considered members of a particular cluster.

- No point in a cluster may be more than 18,000 feet distant (based on right angle routing) from the cluster's centroid.

- No cluster may exceed 1800 lines in size.[27]

- No point in a cluster may be farther than two miles from its nearest neighbor in the cluster.

Note that other than for the wire center boundary restriction, these criteria do not include any arbitrary geographic restrictions, such as clusters being restricted to lie within a single CBG, CB, or latitude/longitude grid cell. Thus, clusters developed pursuant to this process are likely to be the most closely representative of actual telephone distribution areas as determined by outside plant engineers. Note, however, that the last of these restrictions – no point in a cluster may be farther than 2 miles from its nearest neighbor – is not an absolute engineering limitation. It is used to ensure that customer locations that are separated by the given distance are not required to be clustered together. It is possible that efficient engineering of telephone distribution plant would suggest that a different, possibly larger value be chosen.[28]

5.5.2. Clustering Algorithm

The process used by the PNR clustering algorithm is as follows.

a) First, to provide a uniform geographic unit for clustering operations, all customer location geocodes associated with a particular wire center are "rasterized" into 150 foot cells that overlay the geographic rectangle covering the wire center's service area.[29]

b) The algorithm then inspects all neighboring cells (e.g., all cells that have their center within a 150 foot radius of the center of the initial cell) to see if any are populated with customer locations. If one of these neighboring raster cells is populated, the algorithm first checks to see whether clustering that cell (and its immediately surrounding neighbors) with the initial cell would violate any of the

[27] This restriction is based on the maximum unconcentrated lines capacity of an OC-3 fiber optic transmission system used to feed a DLC remote terminal (adjusted for a 90% rate of fill). This is consistent with current HM 5.0 practice that provides a separate channel for each line served by a DLC system. Because it is reasonable engineering practice to concentrate traffic on these large fiber optic DLC systems, future versions of the model may assume that traffic is concentrated on the fiber optic systems feeding DLC remote terminals. When such revisions to the HM become available, the customer location data will be reclustered with the appropriately enlarged maximum limit on cluster size.

[28] Testing of different parameterizations for the maximum distance to a cluster point's nearest neighbor suggests that 2 miles is a reasonable national value.

[29] Rasterization at this level is a process whereby all customer locations that are within the 150 foot square cell are counted as being placed at the center of the cell. Rasterization to 150 foot square cells ensures that the mathematical clustering process can proceed efficiently, and that nonmeaningful distinctions in customer location are ignored. Increasing the raster to 300 or 500 feet would speed up the clustering process dramatically, and still provide a very acceptable level of precision in cluster formation. Note, too, that the PNR calculation of 150 foot raster cells is precise – based on the local latitude of the cell. Thus, raster cells do not enlarge as one moves south toward the equator.

clustering restrictions (i.e., create a cluster that has points more than 18,000 feet from the cluster centroid, create a cluster of more than 1800 lines, or include a point more than 2 miles from its nearest neighbor).[30] If none of these conditions would be violated, the adjoining cell plus its immediate neighbors will be added to the initial cell's cluster

c) This process continues on to the next unclustered populated cell and performs the analysis described in step (b), above. This repeats until no more unclustered cells exist.

d) The process then restarts at step (b), but uses a 300 foot search for populated neighboring cells.

e) This continual enlargement of the search for neighboring populated cells continues on until no more cells can be added to the cluster without violating one or more of the engineering requirements. At this point, the algorithm is complete.

Clusters that contain five or more customer locations are classified as "main" clusters. Clusters that contain from one to four customer locations are called "outlier" clusters. Outlier clusters may be linked to their "home" main cluster via "chains" that string together other outlier clusters that home on the same main cluster. The process of determining the routing of the chain is as follows.

a) The clustering algorithm first calculates the distance between each cluster and every other cluster in the wire center service area.

b) The algorithm then determines the shortest distance between any two clusters, and associates these two clusters together.

c) Next, the algorithm determines the next shortest distance between any two clusters or cluster chains, and associates these together.

d) This process continues until all outliers are chained to a main cluster.

In addition to creating chains to associate outlier clusters to a main cluster, the clustering algorithm calculates the area of the rectangle covering the convex hull of each cluster and the "aspect ratio" of this rectangle.[31]

When this process is completed, the main cluster and its subtending outliers are considered to constitute one serving area.

[30] Because the rasterization into 150 foot square cells may cause customer locations that actually are in the farthest corner of a cell to be considered at the cell's center, the clustering algorithm will actually check to ensure that no cells added to a cluster exceed 17,700 (= 18,000 - 2*150) feet from the cluster's centroid.

[31] The aspect ratio is the ratio of the North-South length to the East-West length of the rectangle covering the convex hull of the cluster.

The description the HM 5.0 Distribution Module in Section 6.3 provides greater detail on the model's engineering of outside plant to serve main and outlier clusters.

5.6. PointCode Translation Processes

PointCode is a Microsoft Access '97 database process that translates between coordinate systems, computes distances and assigns additional characteristics to cluster records. Among the activities executed by PointCode are the following.

Convert the latitude and longitude coordinates provided by PNR for cluster centroids to V&H coordinates. Calculate radial distances between main clusters and their serving wire center. Calculate radial distances between outlier clusters and main clusters.

Compute omega angles between main feeders and the clusters they serve and compute alpha angles between clusters and their subfeeders. Calculate rectilinear (right angle) distances between main clusters and their serving wire center, and between outlier clusters and main clusters.

On the basis of the characteristics of the covering CBG, assign terrain and lines density zone characteristics to the cluster.

6. Module Descriptions

6.1. Input Data Files

6.1.1. Demographic and Geological Parameters

Demographic and geological parameters are obtained from a database developed by PNR and Associates of Jenkintown, PA. Section 5, above, explains in detail how these data are developed. The data file resulting from these processes is organized by state and telephone company (study area). Each customer location cluster (both main and outlier) identified in the wire center service areas of the LEC modeled appears as a separate record in this Microsoft Access 97 database. Each of these cluster records contains the following information:

- Identity of the LEC and wire center serving the cluster;
- Locational information about the cluster relative to its wire center, the predominant CBG in which it falls, its nearest neighboring cluster and associated distances;
- Area and dimensional measurements of the cluster and its lines density;
- Terrain and geological parameters;
- Number of telephone lines by type;

- Number of households and number and type of housing units;
- Number of business firms and employees.

The complete list of data fields in the Cluster Input data table is as follows:

Cluster Input Data Table		
State	Total Area	Households
CLLI	Aspect Ratio	1-HU detach
Company	Company State	1-HU attach
Neca_ID	Density Lines/SQ Mile	HU-2
Group	Rock Depth	HU-4
CBG	Rock Hard	HU-5-9
Cluster Group	Surf Text	HU-10-19
Overall Quad	Water Depth	HU-20-49
Overall Omega	Total Lines	HU-50+
Overall Alpha	Total Bus Lines	Mobile
Radial Dist Feet	Total Res Lines	Other
Cluster or Outlier Check	Special Lines	Firms
Outlier Quad	Public Lines	Employees
Outlier Omega	Single Line Business	
Outlier Alpha		
Outlier Radial Distance		

6.1.2. Wire Center Locations and Interoffice Distances

Calculations to determine total route-miles of interoffice facilities require as inputs several distance measures. These include the distances between each LEC EO and the tandem switch that is assumed to serve it, the distance between the EO and the STP pair that serves it, distances between STPs, distances between tandem offices, and the V&H (vertical and horizontal) coordinates of each switching entity. These data are calculated from a database licensed from Bellcore, referred to as the Special LERG Extract Data ("SLED") file which contains information from the Local Exchange Routing Guide ("LERG"). The SLED includes the V&H coordinates of each switching entity, and the nature of the entity, e.g., end office, tandem, STP, multiple use, etc. Appendix D provides an outline of the process used to develop these distance measures.

6.1.3. ARMIS Data

These data are obtained from the 1996 ARMIS 43-08 Operating Data Reports. ARMIS data are submitted to the FCC annually by all Tier 1 LECs.[32] The following elements of these data are extracted.

- The number of residential access lines, including all residential switched access lines, including those with flat rate (1FR) and measured rate (1MR) service.
- The number of business access lines, including analog single business lines, analog multi-line business lines and digital business lines; these totals include flat rate business (1FB) and measured rate business (1MB) single lines, PBX trunks, Centrex lines, hotel/motel, long distance trunks and multi-line semi-public lines.[33]
- Analog and digital special access lines, including dedicated lines connecting end users' premises to an IXC POP, but do not include intraLATA private lines.
- Public access lines, which include lines associated with coin (public and semi-public) phones, but exclude customer owned pay telephone lines.[34]

For companies that do not report ARMIS, HM 5.0 makes use of data reported in various sources listed earlier in Section 5.1.

6.1.4. User Inputs

This category comprises over 1400 user-definable values. These range from the price of network components, to the percentage of joint-use end offices and tandem offices to capital structure. HAI has supplied default values for each of these variables based on its collective judgment, as augmented by subject matter experts in various areas of network technology, operations and economics. Users can vary these default parameters to reflect unusual local conditions. Appendix B contains a complete description of these variables, along with the default values that have been assigned to them.

6.2. Outside Plant Engineering

The Distribution and Feeder Modules are the main modules controlling the engineering of outside plant. While Sections 6.3 and 6.4 below discuss the unique aspects of each of these modules, this section describes several features and assumptions common to both modules.

---

[32] See, Reporting Requirements for Certain Class A and Tier 1 Telephone Companies (Parts 31, 43, 67 and 69 of the FCC's Rules), CC Docket No. 86-182, 2 FCC Rcd 5770 (1987) (ARMIS Order), modified on recon., 3 FCC Rcd, 6375 (1988). Tier 1 LECs are those with more than $100 million in annual revenues from regulated services. This includes over 50 carriers.

[33] *Id.* at 1-2.

[34] *Id.* at 2.

Figure 5 shows the basic outside plant serving configuration used by HM 5.0.

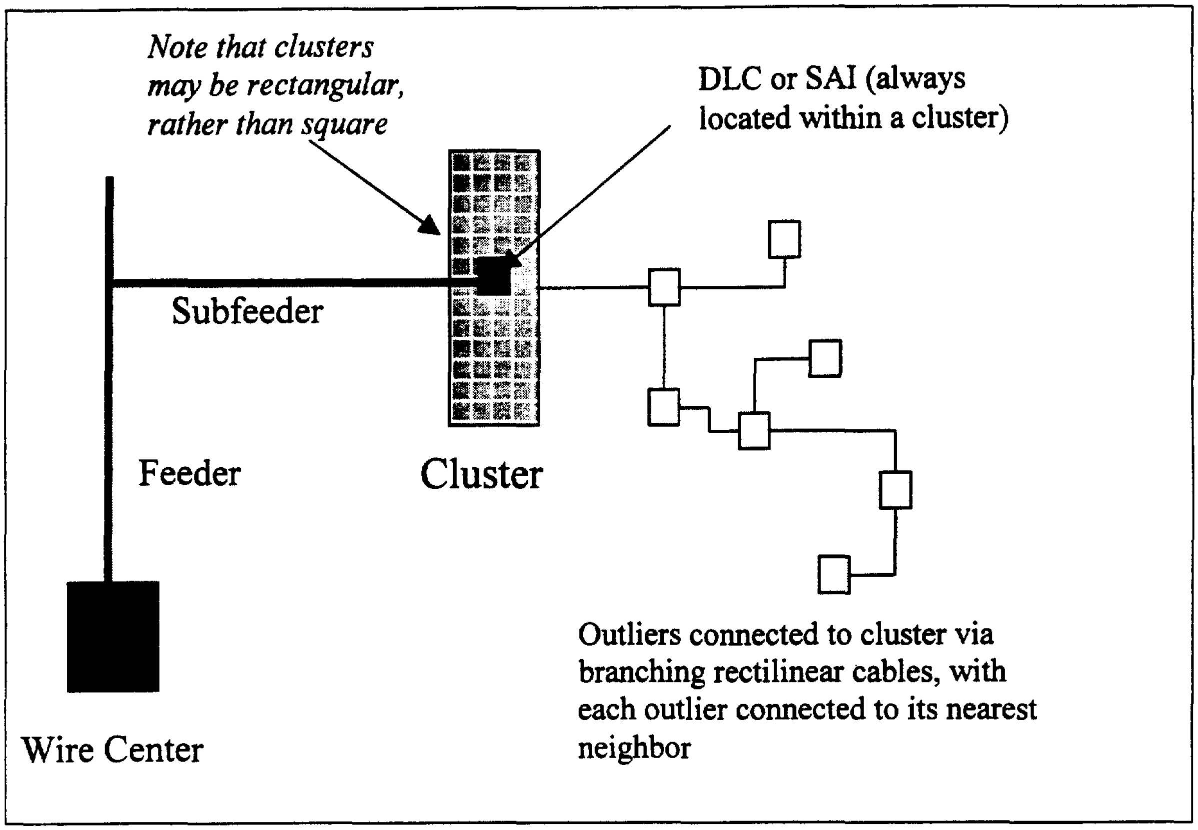

Figure 5 Loop Outside Plant Configuration

The Model assumes a main cluster and its subtending outlier clusters constitute a serving area. Main feeder cable (that may be shared with other main clusters) and subfeeder cable extend from the wire center to an SAI, or, for fiber feeder, a combination of a DLC RT and SAI, at the centroid of the main cluster. The choice between copper and fiber main feeder and subfeeder is made according to the criteria discussed in Section 6.3.5, below. From the SAI, analog copper distribution cables extend in a backbone and branch fashion to all customer locations within the main cluster. The data process used to locate customers and identify population clusters also determines the "aspect ratio" of the overlying rectangle that defines the boundary of a main cluster.

From the centroid of a main cluster, copper cables extend to each outlier cluster that is served from that main cluster, with each outlier cluster connected to its nearest neighbor (either the main cluster or another outlier cluster), via a right-angle route. These copper cables terminate either at an SAI or T1 remote terminal at the centroid of the outlier cluster – depending on whether the distance the signal needs to be carried falls short of, or exceeds, a user-adjustable 18,000 foot threshold. Subscribers in outlier clusters are assumed to be located on routes within the outlier cluster that may be distinct from the route traveled by the cable feeding the outlier's SAI or T1 remote terminal from the main cluster. Because of this, a separate analog copper distribution cable is run from the centroid of the outlier to individual customer locations. The model does, however, assume a moderate amount of structure sharing between these two cables within the outlier cluster because of the partial coincidence of their routes.

6.2.1. Outside Plant Structures

Outside plant structure refers to the set of facilities that support, house, guide, or otherwise protect distribution and feeder cable. There are three types of structure: aerial, buried and underground.

1) *Aerial Structure*

Aerial structure typically consists of poles.[35] Pole investment is a function of the material and labor costs of placing a pole. A user-adjustable input allows the customization of the labor component of pole investment to local conditions. HM 5.0 computes the total investment in aerial distribution and feeder structure within a CBG by evaluating relevant parameters, including the distance between poles, the investment in the pole itself, the total cable sheath mileage, and the fraction of aerial structure along the route.

The model assumes forty-foot Class 4 poles. The spacing between poles for aerial cable is fixed within a given density range but may vary between density ranges. The number of poles on a given route is calculated as:

1 + (route distance/pole spacing), rounded up.

2) *Buried Structure*

Buried structure consists of trenches and related protection against water and other intrusions. The additional cost for protective sheathing and waterproof filling of buried cable is a fixed amount per foot in the case of fiber cable and is a multiplier of cable cost in the case of copper cable.[36] The total investment in buried structure is a function of total route mileage, the fraction of buried structure, investment in protective sheathing and filling, and the density-range-specific cost of trenching.

3) *Underground Structure*

Underground structure consists of conduit and, for feeder plant, manholes and pullboxes. Manholes are used in conjunction with copper cable routes; pullboxes

[35] In the two highest density zones, most aerial structure is assumed to be intrabuilding riser cable and "block cable" attached to buildings.

[36] The default values for sheathing are $0.20 per foot for fiber, and a multiplier of 1.04 for copper. The different treatment reflects the fact that the outside diameter of fiber cable is essentially constant for different strand numbers, while copper cable diameter increases with the number of pairs it contains.